



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: HUSE, WILLIAM D.

(ii) TITLE OF INVENTION: SURFACE EXPRESSION LIBRARIES OF
HETEROMERIC RECEPTORS

(iii) NUMBER OF SEQUENCES: 76

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: CAMPBELL & FLORES LLP
- (B) STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
- (C) CITY: SAN DIEGO
- (D) STATE: CALIFORNIA
- (E) COUNTRY: UNITED STATES
- (F) ZIP: 92122

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/471,622
- (B) FILING DATE: June 5, 1995
- (C) CLASSIFICATION:

(vii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: CAMPBELL, CATHRYN A.
- (B) REGISTRATION NUMBER: 31,815
- (C) REFERENCE/DOCKET NUMBER: P-IX 1613

(viii) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 619-535-9001
- (B) TELEFAX: 619-535-8949

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7445 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATAGCTACTA CTATTAGTAG AATTGATGCC ACCTTTCAG CTCGCGCCCC AAATGAAAAT

60

ATAGCTAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT

120

CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA 180
 GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAAG CTCTAAGCCA 240
 TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300
 TTGGAGTTG CTTCCGGTCT GGTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360
 TCTTCGGGC TTCCTCTAA TCTTTTGAT GCAATCCGCT TTGCTCTGA CTATAATAGT 420
 CAGGGTAAAG ACCTGATTT TGATTTATGG TCATTCTCGT TTTCTGAAC GTTAAAGCA 480
 TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540
 AACATTTA CTATTACCCC CTCTGGCAA ACTTCTTTG CAAAAGCCTC TCGCTATT 600
 GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660
 AATTCCCTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCTAA ATCTCAACTG 720
 ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTT 780
 TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAAATTCA 840
 CAATGATTAA AGTGAAATT AAACCACATCTC AAGCCCAATT TACTACTCGT TCTGGTGT 900
 CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTG TTACGTTGAT TTGGGTAATG 960
 AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020
 TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080
 GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTGCG CGGATTCGA CACAATT 1140
 CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTGCAGC TTGGTATAAT CGCTGGGGGT 1200
 CAAAGATGAG TGTTTAGTG TATTCTTCG CCTCTTCGT TTTAGGTTGG TGCTTCGTA 1260
 GTGGCATTAC GTATTTTACCG CGTTTAATGG AAACCTCCTC ATGAAAAAGT CTTTAGTCCT 1320
 CAAAGCCTCT GTAGCCGTTG CTACCCCTCGT TCCGATGCTG TCTTCGCTG CTGAGGGTGA 1380
 CGATCCCGCA AAAGCGGCCT TAAACTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA 1440
 TGCCTGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTAAGAA 1500
 ATTACACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTT 1560
 TTTTGAGA TTTCAACGT GAAAAAATTA TTATTCGAA TTCTTTAGT TGTTCTTTC 1620
 TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTAGCAA AACCCCATAC AGAAAATTCA 1680
 TTTACTAACG TCTGGAAAGA CGACAAAAGT TTAGATCGTT ACGCTAACTA TGAGGGTTGT 1740
 CTGTGGAATG CTACAGGCGT TGTAGTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA 1800

C3
cont.

TGGGTTCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAACACCTC CTGAGTACGG TGATAACACCT	1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCCTCAGC CTCTTAATAC TTTCATGTT	2040
CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACGT TTTATACGGG CACTGTTACT	2100
CAAGGCAGTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCCTAACCC TCCTGTCAAT	2280
GCTGGCGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT	2400
GATTTGATT ATGAAAAGAT GGCAAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACATTGATT CTGTCGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT	2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCGTGGTG TCTTGCGTT TCTTTATAT GTGCCACCT TTATGTATGT ATTTCTACG	2820
TTTGCTAACAA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTG GGTATTCCGT	2880
TATTATTGCG TTCCCTCGGT TTCCCTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTCATT GTTTCTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTAA CCCTCTGACT	3060
TTGTTCAAGG TGTTCAAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTG	3120
TCTCTGTAAA GGCTGCTATT TTCATTTTG ACGTTAAACA AAAAATCGTT TCTTATTG	3180
ATTGGGATAAA ATAATATGGC TGTTTATTTT GTAACGGCA AATTAGGCTC TGGAAAGACG	3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT	3300
CTTGATTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAC GCCTCGCGTT	3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTGCGTT CTATTGGCG CGGTAATGAT	3420
TCCTACGATG AAAATAAAA CGGCTTGCGTT GTTCTCGATG AGTGCAGGTAC TTGGTTAAT	3480

C3
Cont.

ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTCT ACATGCTCGT 3540
 AAATTAGGAT GGGATATTAT TTTTCTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG 3600
 CGTTCTGCAT TAGCTGAACA TGTTGTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT 3660
 TTTGTCGGTA CTTTATATTTC TCTTATTACT GGCTCGAAAA TGCCCTGCGC TAAATTACAT 3720
 GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT 3780
 ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAACACAGG CTTTTCTAG TAATTATGAT 3840
 TCCGGTGTTC ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA 3900
 AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTGAA AAAAGTTTC ACGCGTTCTT 3960
 TGTCTTGCAG TTGGATTGC ATCAGCATT ACATATAGTT ATATAACCCA ACCTAAGCCG 4020
 GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTGATA AATTCACTAT TGACTCTTCT 4080
 CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT 4140
 AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTATG TACTGTTCC 4200
 ATTAAAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTGTTT TCTTGATGTT 4260
 TGTTTCATCA TCTTCTTTG CTCAGGTAAT TGAAATGAAT AATTGCCCTC TGCGCGATTT 4320
 TGTAACCTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTTCTCCCG ATGTAAGG 4380
 TACTGTTACT GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTT 4440
 TGTTTACGT GCTAATAATT TTGATATGGT TGTTCAATT CCTTCCATAA TTCAGAAGTA 4500
 TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA 4560
 TGATAATTCC GCTCCTCTG GTGGTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC 4620
 TTTTAAAATT AATAACGTTG GGGCAAAGGA TTTAATACGA GTTGTGAAT TGTTGTTAAA 4680
 GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAAC TATTAGTTGT 4740
 TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATT CTTCTACTG TTGATTGCC 4800
 AACTGACCAAG ATATTGATTG AGGGTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA 4860
 TTTTCACTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG 4920
 CCTCACCTCT GTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTAATG GCGATGTTTT 4980
 AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG 5040
 TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTAT 5100
 TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTCAGA CGATTGAGCG 5160

C3
cont.

TCAAAATGTA GGTATTTCCA TGAGCGTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT 5220
 TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT 5280
 TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTACT 5340
 CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA 5400
 AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT 5460
 ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG 5520
 GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT 5580
 TCGCTTCTT CCCTTCCTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC 5640
 GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG 5700
 ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTT CGCCCTTTGA 5760
 CGTTGGAGTC CACGTTCTT AATAGTGGAC TCTTGTCCA AACTGGAACA ACACCTCAACC 5820
 CTATCTCGGG CTATTCTTT GATTATAAG GGATTTGCC GATTTCGGAA CCACCATCAA 5880
 ACAGGATTTT CGCCTGCTGG GGCAAACCAAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG 5940
 CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCC 6000
 GGCGCCCAAT ACGAAACCG CCTCTCCCCG CGCGTTGGCC GATTCAATTAA TGCAGCTGGC 6060
 ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAAT GTGAGTTAGC 6120
 TCACTCATTA GGCACCCCAG GCTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA 6180
 TTGTGAGCGG ATAACAATT CACACGCGTC ACTTGGCACT GGCCGTCGTT TTACAACGTC 6240
 GTGACTGGGA AAACCCCTGGC GTTACCCAAAG CTTTGTACAT GGAGAAAATA AAGTGAAACA 6300
 AAGCACTATT GCACTGGCAC TCTTACCGTT ACCGTTACTG TTTACCCCTG TGACAAAAGC 6360
 CGCCCAGGTC CAGCTGCTCG AGTCAGGCCT ATTGTGCCA GGGGATTGTA CTAGTGGATC 6420
 CTAGGCTGAA GGCAGATGACC CTGCTAACCG TGCATTCAAT AGTTTACAGG CAAGTGCTAC 6480
 TGAGTACATT GGCTACGCTT GGGCTATGGT AGTAGTTATA GTTGGTGCTA CCATAGGGAT 6540
 TAAATTATTC AAAAAGTTA CGAGCAAGGC TTCTTAAGCA ATAGCGAAGA GGCCCGCACC 6600
 GATGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGCGCTTGC CTGGTTCCG 6660
 GCACCAGAAG CGGTGCCGGA AAGCTGGCTG GAGTGCAGTC TTCCCTGAGGC CGATACGGTC 6720
 GTCGTCCCCCT CAAACTGGCA GATGCACGGT TACGATGCGC CCATCTACAC CAACGTAACC 6780
 TATCCCATTA CGGTCAATCC GCCGTTGTT CCCACGGAGA ATCCGACGGG TTGTTACTCG 6840

3
cont.

CTCACATTTA ATGTTGATGA AAGCTGGCTA CAGGAAGGCC AGACGCGAAT TATTTTGAT	6900
GGCGTTCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG AATTAAACA	6960
AAATATTAAC GTTTACAATT TAAATATTTG CTTATACAAT CTTCTGTCTT TTGGGGCTTT	7020
TCTGATTATC AACCGGGGTA CATATGATTG ACATGCTAGT TTTACGATTA CCGTTCATCG	7080
ATTCTCTTGT TTGCTCCAGA CTCTCAGGCA ATGACCTGAT AGCCTTGTA GATCTCTCAA	7140
AAATAGCTAC CCTCTCCGGC ATTAATTAT CAGCTAGAAC GGTTGAATAT CATATTGATG	7200
GTGATTGAC TGTCTCCGGC CTTTCTCACC CTTTGAAATC TTTACCTACA CATTACTCAG	7260
GCATTGCATT TAAAATATAT GAGGGTTCTA AAAATTTTA TCCTTGCCTT GAAATAAAGG	7320
CTTCTCCCGC AAAAGTATTA CAGGGTCATA ATGTTTTGG TACAACCGAT TTAGCTTTAT	7380
GCTCTGAGGC TTTATTGCTT AATTTGCTA ATTCTTGCC TTGCCTGTAT GATTATTGG	7440
ACGTT	7445

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7317 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

C3
Cont

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTCAAG CTCGCGCCCC AAATGAAAAT	60
ATAGCTAAC AGGTTATTGA CCATTGCGA AATGTATCTA ATGGTAAAC TAAATCTACT	120
CGTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAAGATT AGCAATTAAG CTCTAAGCCA	240
TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTTG CTTCCGGTCT GGTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTGAAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTGAT GCAATCCGCT TTGCTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTT TGATTTATGG TCATTCTCGT TTTCTGAACG GTTTAAAGCA	480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT	540
AAACATTTTA CTATTACCCC CTCTGGCAAA ACTCTTTTG CAAAAGCCTC TCGCTATTTT	600
GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT	660

AATTCCTTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCTAA ATCTCAACTG	720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTT	780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA	840
CAATGATTAA AGTGAAATT AAACCATCTC AAGCCAATT TACTACTCGT TCTGGTGT	900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTG TTACGTTGAT TTGGGTAATG	960
AATATCCGGT TCTGTCAAG ATTACTCTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC	1020
TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC	1080
GTCTGCCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTCGA CACAATTTAT	1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT	1200
CAAAGATGAG TGTTTAGTG TATTCTTCG CCTCTTCGT TTTAGGTTGG TGCCTCGTA	1260
GTGGCATTAC GTATTTACC CGTTAACATGG AAACCTCCTC ATGAAAAAGT CTTTAGTCCT	1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTCGCTG CTGAGGGTGA	1380
CGATCCCGCA AAAGCGGCCT TTAACCTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT	1560
TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTGCGAA TTCTTTAGT TGTTCTTTC	1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTAGCAA AACCCCATAC AGAAAATTCA	1680
TTTACTAACG TCTGGAAAGA CGACAAAATC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGTAGTTGT ACTGGTGACG AAAACTCAGTG TTACGGTACA	1800
TGGGTTCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT	2040
CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACGT TTTATACGGG CACTGTTACT	2100
CAAGGCACGT ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTGCG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340

C3
cont.

GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCAGGTCGG GTGGTGGCTC TGTTCCGGT 2400
 GATTTGATT ATGAAAAGAT GGCACACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT 2460
 GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT 2520
 GCTGCTATCG ATGGTTCAT TGGTGACGTT TCCGGCCTG CTAATGGTAA TGTTGCTACT 2580
 GGTGATTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT 2640
 TTAATGAATA ATTCGGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT 2700
 TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTCTATTG ATTGTGACAA AATAAACTTA 2760
 TTCCGTGGTG TCTTGCGTT TCTTTATAT GTGCCACCT TTATGTATGT ATTTCTACG 2820
 TTTGCTAACAA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTG GGTATTCCGT 2880
 TATTATTGCG TTCCCTCGGT TTCCCTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC 2940
 TTAAAAAGGG CTCGGTAAG ATAGCTATTG CTATTCATT GTTTCTTGCT CTTATTATTG 3000
 GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTAA CCCTCTGACT 3060
 TTGTTCAAGGG TGTCAGTTA ATTCTCCGT CTAATGCGCT TCCCTGTTT TATGTTATTC 3120
 TCTCTGTAAA GGCTGCTATT TTCATTTTG ACGTTAAACA AAAAATCGTT TCTTATTG 3180
 ATTGGGATAAA ATAATATGGC TGTTTATTTT GTAATGGCA AATTAGGCTC TGAAAGACCG 3240
 CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT 3300
 CCTGATTAA GGCTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAC GCCTCGCGTT 3360
 CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGCG CGGTAATGAT 3420
 TCCTACGATG AAAATAAAA CGGCTTGCTT GTTCTCGATG AGTGCAGTAC TTGGTTAAT 3480
 ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTCT ACATGCTCGT 3540
 AAATTAGGAT GGGATATTAT TTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG 3600
 CGTTCTGCAT TAGCTGAACA TGTTGTTAT TGTCGTCGTC TGGACAGAAT TACTTACCT 3660
 TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT 3720
 GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGTTTAT 3780
 ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAACACAG CTTTTCTAG TAATTATGAT 3840
 TCCGGTGTGTT ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA 3900
 AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTGA AAAAGTTTC ACGCGTTCTT 3960
 TGTCTTGCAGA TTGGATTGCA ATCAGCATT ACATATAGTT ATATAACCCA ACCTAAGCCG 4020

C3
cont.

GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTGATA AATTCACTAT TGACTCTTCT 4080
 CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAAT 4140
 AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTATG TACTGTTCC 4200
 ATTAAAAAG GTAATTCAA TGAAATTGTT AAATGTAATT AATTTGTTT TCTTGATGTT 4260
 TGTTTCATCA TCTTCTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCCTC TGCGCGATTT 4320
 TGTAACTTGG TATTCAAAGC AATCAGGCAG ATCCGTTATT GTTTCTCCCG ATGTAAAAGG 4380
 TACTGTTACT GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTC 4440
 TGTTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA 4500
 TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA 4560
 TGATAATTCC GCTCCTTCTG GTGGTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC 4620
 TTTTAAAATT AATAACGTTC GGGCAAAGGA TTTAATACGA GTTGTGAAT TGTTTGTAAA 4680
 GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAAC TATTAGTTGT 4740
 TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTTGCC 4800
 AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA 4860
 TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG 4920
 CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTAATG GCGATGTTTT 4980
 AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG 5040
 TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTAT 5100
 TACTGGTCGT GTGACTGGTG AATCTGCCAA TGAAATAAT CCATTCAGA CGATTGAGCG 5160
 TCAAAATGTA GGTATTTCCA TGAGCGTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT 5220
 TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT 5280
 TACTAATCAA AGAAGTATTG CTACAACGGT TAATTGCGT GATGGACAGA CTCTTTACT 5340
 CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA 5400
 AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT 5460
 ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTAA AGCGCGCGG 5520
 GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT 5580
 TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC 5640
 GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCCC AAAAAACTTG 5700

C³
cont.

ATTTGGGTGA TGGTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTT CGCCCTTGAA 5760
 CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCGA AACTGGAACA ACACTCAACC 5820
 CTATCTCGGG CTATTCTTT GATTTATAAG GGATTTGCC GATTTCGGAA CCACCATCAA 5880
 ACAGGATTT CGCCTGCTGG GGCAAACCAAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG 5940
 CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCC 6000
 GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGCC GATTCATTA TGCAAGCTGGC 6060
 ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC 6120
 TCACTCATTA GGCACCCCAG GCTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA 6180
 TTGTGAGCGG ATAACAATT CACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC 6240
 TACGGCAGCC GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCG AGCTCGTGAT 6300
 GACCCAGACT CCAGATATCC AACAGGAATG AGTGTAAATT CTAGAACGCG TCACTTGGCA 6360
 CTGGCCGTCG TTTTACAACG TCGTGAUTGG GAAAACCTG GCGTTACCCA AGCTTAATCG 6420
 CCTTGCAGAA TTCCCTTCG CCAGCTGGCG TAATAGCGAA GAGGCCGCA CCGATCGCCC 6480
 TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTT GCCTGGTTTC CGGCACCAGA 6540
 AGCGGTGCCG GAAAGCTGGC TGGAGTGCAG TCTTCCTGAG GCCGATAACGG TCGTCGTCCC 6600
 CTCAAACTGG CAGATGCACG GTTACGATGC GCCCATCTAC ACCAACGTAAC CCTATCCCAT 6660
 TACGGTCAAT CCGCCGTTTG TTCCCACGGA GAATCCGACG GGTTGTTACT CGCTCACATT 6720
 TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACGCGA ATTATTTTG ATGGCGTTCC 6780
 TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTAA CAAAATATTA 6840
 ACGTTACAA TTTAAATATT TGCTTATACA ATCTTCCTGT TTTGGGGCT TTTCTGATTA 6900
 TCAACCGGGG TACATATGAT TGACATGCTA GTTTACGAT TACCGTTCAT CGATTCTCTT 6960
 GTTTGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTG TAGATCTCTC AAAAATAGCT 7020
 ACCCTCTCCG GCATTAATTT ATCAGCTAGA ACGGTTGAAT ATCATATTGA TGGTGATTTG 7080
 ACTGTCTCCG GCCTTCTCA CCCTTTGAA TCTTACCTA CACATTACTC AGGCATTGCA 7140
 TTTAAAATAT ATGAGGGTTC TAAAAATTT TATCCTTGCG TTGAAATAAA GGCTTCTCCC 7200
 GCAAAAGTAT TACAGGGTCA TAATGTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG 7260
 GCTTTATTGC TTAATTTGC TAATTCTTG CCTTGCTGT ATGATTTATT GGATGTT 7317

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7729 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTCA G CTCGCGCCCC AAATGAAAAT 60
 ATAGCTAAC AGGTTATTGA CCATTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120
 CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA 180
 GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTG AGCAATTAAAG CTCTAAGCCA 240
 TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300
 TTGGAGTTG CTTCCGGTCT GGTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360
 TCTTCGGGC TTCCTCTTAA TCTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420
 CAGGGTAAAG ACCTGATTT TGATTTATGG TCATTCTCGT TTTCTGAACG GTTTAAAGCA 480
 TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540
 AACACATTAA CTATTACCCC CTCTGGCAAA ACTTCTTTG CAAAAGCCTC TCGCTATTTT 600
 GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660
 AATTCCCTTT GGCCTTATGT ATCTGCATTA GTTGAATGTG GTATTCTCAA ATCTCAACTG 720
 ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTAA 780
 TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA 840
 CAATGATTAA AGTGAAATT AAACCACCTC AAGCCCAATT TACTACTCGT TCTGGTGTAA 900
 CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTG TTACGTTGAT TTGGGTAATG 960
 AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020
 TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080
 GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTG CGGATTCGA CACAATTAA 1140
 CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTGCAGC TTGGTATAAT CGCTGGGGGT 1200
 CAAAGATGAG TGTTTAGTG TATTCTTCG CCTCTTCGT TTTAGGTTGG TGCCTTCGTA 1260
 GTGGCATTAC GTATTTTAC CGTTTAATGG AAACCTCCTC ATGAAAAAGT CTTAGTCCT 1320
 CAAAGCCTCT GTAGCCGTTG CTACCCCTCGT TCCGATGCTG TCTTCGCTG CTGAGGGTGA 1380

C3
 cont.

CGATCCCGCA AAAGCGGCCT TTAACCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA	1440
TGCGTGGCG ATGGTTGTTG TCATTGTCGG CGCAAATC GGTATCAAGC TGTTTAAGAA	1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTT GGAGCCTTT	1560
TTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTCGAA TTCCTTAGT TGTTCCCTTC	1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTAGCAA AACCCCATAAC AGAAAATTCA	1680
TTTACTAACG TCTGGAAAGA CGACAAAATC TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGCGT TGTAGTTGT ACTGGTGACG AAAACTCAGTG TTACGGTACA	1800
TGGGTTCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCCTGG TACTGAGCAA	1980
AACCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT	2040
CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT	2100
CAAGGCACTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTCAAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGTTCCGGT	2400
GATTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT	2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT	2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAAACTTA	2760
TTCCGTGGTG TCTTTGCGTT TCTTTATAT GTGCCACCT TTATGTATGT ATTTCTACG	2820
TTTGCTAACAA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTG GGTATTCCGT	2880
TATTATTGCG TTTCCCTCGGT TTCCTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTCATT GTTTCTTGCT CTTATTATTG	3000
GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT	3060

3
cont

TTGTTCAAGGG TGTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATT 3120
 TCTCTGTAAA GGCTGCTATT TTCATTTTG ACGTTAAACA AAAAATCGTT TCTTATTG 3180
 ATTGGGATAA ATAATATGGC TGTTTATTGTA GTAAGTGGCA AATTAGGCTC TGAAAGACG 3240
 CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGAAAAT AGCAACTAAT 3300
 CTTGATTAA GGCTCAAAA CCTCCCGCAA GTGGGAGGT TCGCTAAAC GCCTCGCGTT 3360
 CTTAGAACAC CGGATAAGCC TTCTATATCT GATTTGCTG CTATTGGCG CGGTAATGAT 3420
 TCCTACGATG AAAATAAAA CGGCTTGCTT GTTCTCGATG AGTGCAGTAC TTGGTTAAT 3480
 ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTCT ACATGCTCGT 3540
 AAATTAGGAT GGGATATTAT TTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG 3600
 CGTTCTGCAT TAGCTGAACA TGTTGTTAT TGTCGTCGTC TGGACAGAAT TACTTACCT 3660
 TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAA TGCGCTGCC TAAATTACAT 3720
 GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGTTTAT 3780
 ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAACACAGG CTTTTCTAG TAATTATGAT 3840
 TCCGGTGTGTT ATTCTTATTAA AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA 3900
 AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTGAA AAAAGTTTC ACGCGTTCTT 3960
 TGTCTTGCAGA TTGGATTGTC ATCAGCATT ACATATAGTT ATATAACCCA ACCTAAGCCG 4020
 GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTGATA AATTCACTAT TGACTCTCT 4080
 CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGAAA ATTAATTAAT 4140
 AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTATG TACTGTTCC 4200
 ATTAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTGTTT TCTTGATGTT 4260
 TGTTTCATCA TCTTCTTTG CTCAGGTAAT TGAAATGAAT AATTGCCCTC TGCGCGATTT 4320
 TGTAACCTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTCTCCCG ATGTAAGG 4380
 TACTGTTACT GTATATTCACT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTATTTC 4440
 TGTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA 4500
 TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA 4560
 TGATAATTCC GCTCCTCTG GTGGTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC 4620
 TTTTAAAATT AATAACGTTC GGGCAAAGGA TTTAATACGA GTTGTGAAT TGTTGTAAA 4680
 GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT 4740

C3
cont.

TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATT CTTTCTACTG TTGATTTGCC	4800
AACTGACCAG ATATTGATTG AGGGTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA	4860
TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG	4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTAATG GCGATGTTTT	4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTAT	5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAATAAT CCATTCAGA CGATTGAGCG	5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTGCGT GATGGACAGA CTCTTTACT	5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA	5400
AATCCCTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTAA AGCGCGGC	5520
GTGTGGTGGT TACCGCGAGC GTGACCGCTA CACTGCCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCCC AAAAAACTTG	5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA	5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTCCA AACTGGAACA ACACCTCAACC	5820
CTATCTCGGG CTATTCTTTT GATTATAAG GGATTTGCC GATTTCGGAA CCACCATCAA	5880
ACAGGATTT CGCCTGCTGG GGCAAACCAAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCT	6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCAATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAAT GTGAGTTAGC	6120
TCACTCATTA GGCACCCCCAG GCTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTT CACACCGTC ACTTGGCACT GGCGTCGTT TTACAACGTC	6240
GTGACTGGGA AAACCCCTGGC GTTACCCAAG CTTTGTACAT GGAGAAAATA AAGTGAACAA	6300
AAGCACTATT GCACGGCAC TCTTACCGTT ACTGTTTACC CCTGTGGCAA AAGCCCAGGT	6360
CCAGCTGCTC GAGTCGGTCT TCCCCCTGGC ACCCTCCTCC AAGAGCACCT CTGGGGGCAC	6420

C3
cont.

AGCGGCCCTG GGCTGCCTGG TCAAGACTAA TTCCCCAAC CGGTGACGGT GTCGTGGAAC 6480
 TCAGGCGCCC TGACCAGCGG CGTGCACACC TTCCCGGCTG TCCTACAGTC CTCAGGACTC 6540
 TACTCCCTCA GCAGCGTGGT GACCGTGCCC TCCAGCAGCT TGGGCACCCA GACCTACATC 6600
 TGCAACGTGA ATCACAAAGCC CAGCAACACC AAGGTGGACA AGAAAGCAGA GCCCAAATCT 6660
 TGTACTAGTG GATCCTACCC GTACGACGTT CCGGACTACG CTTCTTAGGC TGAAGGCGAT 6720
 GACCCCTGCTA AGGCTGCATT CAATAGTTA CAGGCAAGTG CTACTGAGTA CATTGGCTAC 6780
 GCTTGGGCTA TGGTAGTAGT TATAGTTGGT GCTACCATAG GGATTAAATT ATTCAAAAAG 6840
 TTTACGAGCA AGGCTTCTTA AGCAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC 6900
 AGTTGCGCAG CCTGAATGGC GAATGGCGCT TTGCCTGGTT TCCGGCACCA GAAGCGGTGC 6960
 CGGAAAGCTG GCTGGAGTGC GATCTCCTG AGGCCGATAC GGTCGTCGTC CCCTCAAAC 7020
 GGCAGATGCA CGGTTACGAT GCGCCCATCT ACACCAACGT AACCTATCCC ATTACGGTCA 7080
 ATCCGCCGTT TGTTCCCACG GAGAATCCGA CGGGTTGTTA CTCGCTCACA TTTAATGTTG 7140
 ATGAAAGCTG GCTACAGGAA GGCCAGACGC GAATTATTTTG TGATGGCGTT CCTATTGGTT 7200
 C3
 C
 cont
 AAAAATGAG CTGATTTAAC AAAAATTAA CGCGAATTAA AACAAAATAT TAACGTTTAC 7260
 AATTTAAATA TTTGCTTATA CAATCTCCT GTTTTGGGG CTTTTCTGAT TATCAACCAGG 7320
 GGTACATATG ATTGACATGC TAGTTTACG ATTACCGTTC ATCGATTCTC TTGTTGCTC 7380
 CAGACTCTCA GGCAATGACC TGATAGCCTT TGTAGATCTC TCAAAAATAG CTACCCCTCTC 7440
 CGGCATTAAT TTATCAGCTA GAACGGTTGA ATATCATATT GATGGTGATT TGACTGTCTC 7500
 CGGCCTTCT CACCTTTTG AATCTTAC TACACATTAC TCAGGCATTG CATTAAAAT 7560
 ATATGAGGGT TCTAAAAATT TTTATCCTTG CGTTGAAATA AAGGCTCTC CCGCAAAAGT 7620
 ATTACAGGGT CATAATGTTT TTGGTACAAC CGATTTAGCT TTATGCTCTG AGGCTTTATT 7680
 GCTTAATTTC GCTAATTCTT TGCCTTGCCT GTATGATTAA TTGGACGTT 7729

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 7557 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTCA G CTCGCCCCC AAATGAAAAT 60
 ATAGCTAAAC AGTTTATTGA CCATTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120
 CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA 180
 GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTG AGCAATTAAG CTCTAAGCCA 240
 TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300
 TTGGAGTTG CTTCCGGTCT GGTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360
 TCTTCGGGC TTCCTCTTAA TCTTTTGAT GCAATCCGCT TTGCTCTGA CTATAATAGT 420
 CAGGGTAAAG ACCTGATTT TGATTTATGG TCATTCTCGT TTTCTGAAC GTTAAAGCA 480
 TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540
 AACACATTTA CTATTACCCC CTCTGGCAAA ACTTCTTTG CAAAAGCCTC TCGCTATTTT 600
 GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660
 AATTCCCTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCTAA ATCTCAACTG 720
 ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT 780
 TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA 840
 CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGT 900
 CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTG TTACGTTGAT TTGGGTAATG 960
 AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020
 TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080
 GTCTGCGCCT CGTCCGGCT AAGTAACATG GAGCAGGTG CGGATTTCGA CACAATTAT 1140
 CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT 1200
 CAAAGATGAG TGTTTAGTG TATTCTTCG CCTCTTTCGT TTTAGGTTGG TGCCCTCGTA 1260
 GTGGCATTAC GTATTTACC CGTTAACATGG AAACCTCCTC ATGAAAAAGT CTTTAGTCCT 1320
 CAAAGCCTCT GTAGCCGTTG CTACCCCTCGT TCCGATGCTG TCTTCGCTG CTGAGGGTGA 1380
 CGATCCCCGA AAAGCGGCCT TTAACCTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA 1440
 TGCCTGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTAAGAA 1500
 ATTACACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTT GGAGCCTTTT 1560
 TTTTGAGA TTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTAGT TGTTCCCTTC 1620
 TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTAGCAA AACCCCATAC AGAAAATTCA 1680

3
cont.

TTTACTAACG TCTGGAAAGA CGACAAAACT TTAGATCGTT ACGCTAACTA TGAGGGTTGT	1740
CTGTGGAATG CTACAGGGCGT TGTAGTTGT ACTGGTGACG AAAACTCAGTG TTACGGTACA	1800
TGGGTTCTCA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT	1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT	1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA	1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT	2040
CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT	2100
CAAGGCACTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG	2160
TATGACGCTT ACTGGAACGG TAAATTAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA	2220
GATCCATTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT	2280
GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT	2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGG GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT	2400
GATTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT	2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT	2520
<i>C3</i> <i>cont.</i>	
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT	2580
GGTGATTGGT CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT	2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGTTGA ATGTCGCCCT	2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA	2760
TTCCGTGGTG TCTTTGCGTT TCTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG	2820
TTTGCTAAC A TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTG GGTATTCCGT	2880
TATTATTGCG TTCCCTCGGT TTCCCTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC	2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CCTGTTCTT GCTCTTATTA TTGGGCTTAA	3000
CTCAATTCTT GTGGGTTATC TCTCTGATAT TAGCGCTCAA TTACCCCTCG ACTTTGTTCA	3060
GGGTGTTCAAG TTAATTCTCC CGTCTAATGC GCTCCCTGT TTTTATGTTA TTCTCTCTGT	3120
AAAGGCTGCT ATTTCATTT TTGACGTTAA ACAAAAATC GTTTCTTATT TGGATTGGGA	3180
TAAATAATAT GGCTGTTAT TTTGTAACTG GCAAATTAGG CTCTGGAAAG ACGCTCGTTA	3240
GCGTTGGTAA GATTCAAGGAT AAAATTGTAG CTGGGTGCAA AATAGCAACT AATCTTGATT	3300
TAAGGCTTCA AACCTCCCG CAAGTCGGGA GGTCGCTAA AACGCCTCGC GTTCTTAGAA	3360

TACCGGATAA GCCTTCTATA TCTGATTGC TTGCTATTGG GCGCGGTAAT GATTCCCTACG 3420
 ATGAAAATAA AAACGGCTTG CTTGTTCTCG ATGAGTGCAG TACTTGGTT AATACCCGTT 3480
 CTTGGAATGA TAAGGAAAGA CAGCCGATTA TTGATTGGTT TCTACATGCT CGTAAATTAG 3540
 GATGGGATAT TATTTTCCTT GTTCAGGACT TATCTATTGT TGATAAACAG GCGCGTTCTG 3600
 CATTAGCTGA ACATGTTGTT TATTGTCGTC GTCTGGACAG AATTACTTTA CCTTTGTCG 3660
 GTACTTTATA TTCTCTTATT ACTGGCTCGA AAATGCCTCT GCCTAAATTA CATGTTGGCG 3720
 TTGTTAAATA TGGCGATTCT CAATTAAGCC CTACTGTTGA GCGTTGGCTT TATACTGGTA 3780
 AGAATTGTA TAACGCATAT GATACTAAAC AGGCTTTTC TAGTAATTAT GATTCCGGTG 3840
 TTTATTCTTA TTTAACGCCT TATTTATCAC ACGGTCGGTA TTTCAAACCA TTAAATTTAG 3900
 GTCAGAAGAT GAAGCTTACT AAAATATATT TGAAAAAGTT TTCACGCGTT CTTTGTCTTG 3960
 CGATTGGATT TGCATCAGCA TTTACATATA GTTATATAAC CCAACCTAAG CCGGAGGTTA 4020
 AAAAGGTAGT CTCTCAGACC TATGATTTC TATAAATTAC TATTGACTCT TCTCAGCGTC 4080
 TTAATCTAAG CTATCGCTAT GTTTCAAGG ATTCTAAGGG AAAATTAATT AATAGCGACG 4140
 ATTTACAGAA GCAAGGTTAT TCACTCACAT ATATTGATTT ATGTACTGTT TCCATTAAAA 4200
 C
 cont.
 AAGGTAATTC AAATGAAATT GTTAAATGTA ATTAATTGGT TTTTCTTGAT GTTTGTTTCA 4260
 TCATCTTCTT TTGCTCAGGT AATTGAAATG AATAATTGCA CTCTGCGCGA TTTTGTAACT 4320
 TGGTATTCAA AGCAATCAGG CGAACCGTT ATTGTTCTC CCGATGTAAA AGGTACTGTT 4380
 ACTGTATATT CATCTGACGT TAAACCTGAA AATCTACGCA ATTTCTTAT TTCTGTTTTA 4440
 CGTGCTAATA ATTTGATAT GGTTGGTCA ATTCCTTCCA TAATTCAAGAA GTATAATCCA 4500
 AACAAATCAGG ATTATATTGA TGAATTGCCA TCATCTGATA ATCAGGAATA TGATGATAAT 4560
 TCCGCTCCTT CTGGTGGTTT CTTTGTCCG CAAAATGATA ATGTTACTCA AACTTTAAA 4620
 ATTAATAACG TTCGGGCAAA GGATTTAATA CGAGTTGTCG AATTGTTGT AAAGTCTAAT 4680
 ACTTCTAAAT CCTCAAATGT ATTATCTATT GACGGCTCTA ATCTATTAGT TGTTAGTGCA 4740
 CCTAAAGATA TTTTAGATAA CCTTCCTCAA TTCCCTTCTA CTGTTGATTT GCCAACTGAC 4800
 CAGATATTGA TTGAGGGTTT GATATTGAG GTTCAGCAAG GTGATGCTTT AGATTTTCA 4860
 TTTGCTGCTG GCTCTCAGCG TGGCACTGTT GCAGGGCGGTG TTAATACTGA CCGCCTCACC 4920
 TCTGTTTAT CTTCTGCTGG TGGTTCGTTC GGTATTTTA ATGGCGATGT TTTAGGGCTA 4980
 TCAGTTCGCG CATTAAAGAC TAATAGCCAT TCAAAATAT TGTCTGTGCC ACGTATTCTT 5040

ACGCTTCAG GTCAGAAGGG TTCTATCTCT GTTGGCCAGA ATGTCCCTTT TATTACTGGT	5100
CGTGTGACTG GTGAATCTGC CAATGTAAT AATCCATTTC AGACGATTGA GCGTCAAAAT	5160
GTAGGTATTT CCATGAGCGT TTTCCCTGTT GCAATGGCTG GCGGTAATAT TGTTCTGGAT	5220
ATTACCAGCA AGGCCGATAG TTTGAGTTCT TCTACTCAGG CAAAGTGATGT TATTACTAAT	5280
CAAAGAAGTA TTGCTACAAC GGTTAATTG CGTGATGGAC AGACTCTTT ACTCGGTGGC	5340
CTCACTGATT ATAAAAACAC TTCTCAAGAT TCTGGCGTAC CGTTCCGTGTC TAAAATCCCT	5400
TTAACCGGCC TCCTGTTAG CTCCCGCTCT GATTCCAACG AGGAAAGCAC GTTATACGTG	5460
CTCGTCAAAG CAACCATAGT ACGCGCCCTG TAGCGCGCA TTAAGCGCGG CGGGTGTGGT	5520
GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT	5580
CTTCCCTTCC TTTCTCGCCA CGTTGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT	5640
CCCTTTAGGG TTCCGATTTA GTGCTTACG GCACCTCGAC CCCAAAAAAC TTGATTTGGG	5700
TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTGCCCTT TGACGTTGGA	5760
GTCCACGTTC TTTAATAGTG GACTCTGTT CCAAACCTGGA ACAACACTCA ACCCTATCTC	5820
GGGCTATTCT TTTGATTTAT AAGGGATTTT GCCGATTCG GAACCACCAT CAAACAGGAT	5880
TTTCGCCTGC TGGGGCAAAC CAGCGTGGAC CGCTTGCTGC AACTCTCTCA GGGCCAGGCG	5940
GTGAAGGGCA ATCAGCTGTT GCCCGTCTCG CTGGTGAAAA GAAAAACAC CCTGGCGCCC	6000
AATACGCAA CCCCTCTCC CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG	6060
GTTTCCGAC TGGAAAGCGG GCAGTGAGCG CAACGCAATT AATGTGAGTT AGCTCACTCA	6120
TTAGGCACCC CAGGCTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG	6180
CGGATAACAA TTTCACACGC CAAGGAGACA GTCATAATGA AATACCTATT GCCTACGGCA	6240
GCCGCTGGAT TGTATTACT CGCTGCCAA CCAGCCATGG CCGAGCTCTT CCCGCCATCT	6300
GATGAGCAGT TGAAATCTGG AACTGCCCTCT GTTGTGTGCC TGCTGAATAA CTTCTATCCC	6360
AGAGAGGCCA AAGTACAGTG GAAGGTGGAT AACGCCCTCC AATCGGGTAA CTCCCAAGGAG	6420
AGTGTACACAG AGCAGGACAG CAAGGACAGC ACCTACAGCC TCAGCAGCAC CCTGACGCTG	6480
AGCAAAGCAG ACTACGAGAA ACACAAAGTC TACGCCTGCG AAGTCACCCA TCAGGGCCTG	6540
AGCTCGCCCG TCACAAAGAG CTTCAACAGG GGAGAGTGTT CTAGAACGCG TCACCTGGCA	6600
CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCTG GCGTTACCCA AGCTTAATCG	6660
CCTTGCAGAA TTCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC	6720

C³
cont.

TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTT GCCTGGTTTC CGGCACCAGA	6780
AGCGGTGCCG GAAAGCTGGC TGGAGTGCAG TCTTCCTGAG GCCGATACGG TCGTCGTCCC	6840
CTCAAACCTGG CAGATGCACG GTTACGATGC GCCCATCTAC ACCAACGTAAC CCTATCCCAT	6900
TACGGTCAAT CCGCCGTTG TTCCCACGGA GAATCCGACG GGTTGTTACT CGCTCACATT	6960
TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACGCGA ATTATTTTG ATGGCGTTCC	7020
TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTAA CAAAATATTA	7080
ACGTTACAA TTTAAATATT TGCTTATACA ATCTTCCTGT TTTTGGGCT TTTCTGATTA	7140
TCAACCGGGG TACATATGAT TGACATGCTA GTTTACGAT TACCGTTCAT CGATTCTCTT	7200
GTTTGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTG TAGATCTCTC AAAAATAGCT	7260
ACCCCTCTCCG GCATTAATTT ATCAGCTAGA ACGGTTGAAT ATCATATTGA TGGTGATTTG	7320
ACTGTCTCCG GCCTTCTCA CCCTTTGAA TCTTTACCTA CACATTACTC AGGCATTGCA	7380
TTTAAAATAT ATGAGGGTTC TAAAAATTTT TATCCTTGCG TTGAAATAAA GGCTTCTCCC	7440
GCAAAAGTAT TACAGGGTCA TAATGTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG	7500
GCTTTATTGC TTAATTTGC TAATTCTTG CCTTGCTGT ATGATTATT GGATGTT	7557

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8118 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTCAAG CTCGCGCCCC AAATGAAAAT	60
ATAGCTAAAC AGGTTATTGA CCATTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT	120
CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA	180
GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAAG CTCTAAGCCA	240
TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG	300
TTGGAGTTG CTTCCGGTCT GGTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG	360
TCTTTCGGGC TTCCTCTTAA TCTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT	420
CAGGGTAAAG ACCTGATTG TGATTATGG TCATTCTCGT TTTCTGAACG GTTTAAAGCA	480

TTTGAGGGGG ATTCAATGAA TATTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540
 AAACATTTA CTATTACCCC CTCTGGAAA ACTTCTTTG CAAAAGCCTC TCGCTATT 600
 GGTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCCTCGT 660
 AATTCCCTTT GGC GTTATGT ATCTGCATTA GTGAATGTG GTATTCTAA ATCTCAACTG 720
 ATGAATCTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATT 780
 TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA 840
 CAATGATTAA AGTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGT 900
 CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTG TTACGTTGAT TTGGGTAATG 960
 AATATCCGGT TCTTGTCAAG ATTACTCTTGT ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020
 TGTACACCGT TCATCTGTCC TCTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080
 GTCTGCGCCT CGTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATT 1140
 CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGT 1200
 CAAAGATGAG TGTTTAGTG TATTCTTCG CCTCTTCGT TTTAGGTTGG TGCCCTCGTA 1260
 GTGGCATTAC GTATTTACC CGTTAATGG AAACCTCCTC ATGAAAAAGT CTTTAGTCCT 1320
 CAAAGCCTCT GTAGCCGTTG CTACCCCTCGT TCCGATGCTG TCTTCGCTG CTGAGGGTGA 1380
 CGATCCCGCA AAAGCGGCCT TTAACCTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA 1440
 TCGTGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTAAGAA 1500
 ATTACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTT GGAGCCTTT 1560
 TTTTGGAGA TTTCAACGT GAAAAAATTA TTATTGCAA TTCTTTAGT TGTTCCCTTC 1620
 TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTAGCAA AACCCCATAC AGAAAATTCA 1680
 TTTACTAACG TCTGGAAAGA CGACAAAATC TTAGATCGTT ACGCTAACTA TGAGGGTTGT 1740
 CTGTGGAATG CTACAGGCGT TGTAGTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA 1800
 TGGGTTCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT 1860
 TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT 1920
 ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCTGG TACTGAGCAA 1980
 AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTT 2040
 CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAAC TGTTATACGGG CACTGTTACT 2100
 CAAGGCAGTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG 2160

C3
cont.

TATGACGCTT ACTGGAACGG TAAATTCAAGA GACTGCGCTT TCCATTCTGG CTTAACGAA 2220
 GATCCATTCTG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCGCTAACCTCCTGTCAAT 2280
 GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT 2340
 GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGTTCCGGT 2400
 GATTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT 2460
 GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACATTGATT CTGTCGCTAC TGATTACGGT 2520
 GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGTTGCTACT 2580
 GGTGATTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT 2640
 TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGGTGA ATGTCGCCCT 2700
 TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTCTATTG ATTGTGACAA AATAAACTTA 2760
 TTCCGTGGTG TCTTGCGTT TCTTTATAT GTGCCACCT TTATGTATGT ATTTCTACG 2820
 TTTGCTAACAA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTG GGTATTCCGT 2880
 TATTATTGCG TTTCCTCGGT TTCCTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC 2940
 TAAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTCATT GTTCTTGCT CTTATTATTG 3000
 GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTAA CCCTCTGACT 3060
 TTGTTCAAGGG TGTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTT TATGTTATTC 3120
 TCTCTGTAAA GGCTGCTATT TTCATTTTG ACGTTAAACA AAAAATCGTT TCTTATTG 3180
 ATTGGGATAAA ATAATATGGC TGTTTATTTT GTAACTGGCA AATTAGGCTC TGAAAGACG 3240
 CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGAAAAT AGCAACTAAT 3300
 CTTGATTTAA GGCTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAC GCCTCGCGTT 3360
 CTTAGAATAC CGGATAAGCC TTCTATATCT GATTGCTTG CTATTGGCG CGGTAATGAT 3420
 TCCTACGATG AAAATAAAA CGGCTTGCTT GTTCTCGATG AGTGCAGTAC TTGGTTAAT 3480
 ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTCT ACATGCTCGT 3540
 AAATTAGGAT GGGATATTAT TTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG 3600
 CGTTCTGCAT TAGCTGAACA TGTTGTTAT TGTCGTCGTC TGGACAGAAT TACTTACCT 3660
 TTTGTCGGTA CTTTATATTCT CTTTATTACT GGCTGAAAAA TGCGCTTGCC TAAATTACAT 3720
 GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGTTTAT 3780
 ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTCTAG TAATTATGAT 3840

C3
cont.

TCCGGTGT	TTT ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA	3900
AATTTAGGTC	AGAAGATGAA GCTTACTAAA ATATATTGAA AAAAGTTTC ACGCGTTCTT	3960
TGTCTTGCGA	TTGGATTGTC ATCAGCATT ACATATAGTT ATATAACCCA ACCTAACGCCG	4020
GAGGTTAAA	AGGTAGTCTC TCAGACCTAT GATTTGATA AATTCACTAT TGACTCTCT	4080
CAGCGTCTTA	ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT	4140
AGCGACGATT	TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTCC	4200
ATTAAAAAAAG	GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTGTTT TCTTGATGTT	4260
TGTTTCATCA	TCTTCTTTG CTCAGGTAAT TGAAATGAAT AATTGCCTC TGCGCGATT	4320
TGTAACCTGG	TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTCTCCCG ATGTAAAAGG	4380
TACTGTTACT	GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTC	4440
TGTTTACGT	GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCATCAA TTCAGAAGTA	4500
TAATCCAAAC	AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA	4560
TGATAATTCC	GCTCCTCTG GTGGTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC	4620
TTTTAAAATT	AATAACGTTG GGGCAAAGGA TTTAATACGA GTTGTGAAAT TGTTGTAAA	4680
GTCTAAACT	TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAAC TATTAGTTGT	4740
TAGTGCACCT	AAAGATATT TAGATAACCT TCCTCAATT CTTTCTACTG TTGATTTGCC	4800
AACTGACCAG	ATATTGATTG AGGGTTGAT ATTGAGGTT CAGCAAGGTG ATGCTTACA	4860
TTTTTCATTT	GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG	4920
CCTCACCTCT	GTTTATCTT CTGCTGGTGG TTCGTTGGT ATTTTAATG GCGATGTTT	4980
AGGGCTATCA	GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG	5040
TATTCTTACG	CTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTAT	5100
TACTGGTCGT	GTGACTGGTG AATCTGCCAA TGAAATAAT CCATTCAGA CGATTGAGCG	5160
TCAAAATGTA	GGTATTCCA TGAGCGTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT	5220
TCTGGATATT	ACCAGCAAGG CCGATAGTTT GAGTTCTCT ACTCAGGCAA GTGATGTTAT	5280
TACTAATCAA	AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTACT	5340
CGGTGGCCTC	ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA	5400
AATCCCTTTA	ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT	5460
ATACGTGCTC	GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGGGG	5520

3
cont.

GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTGCCAG CGCCCTAGCG CCCGCTCCTT	5580
TCGCTTCCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC	5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG	5700
ATTGGGGTGA TGGTCACGT AGTGGGCAT CGCCCTGATA GACGGTTTT CGCCCTTGA	5760
CGTTGGAGTC CACGTTCTT AATAGTGGAC TCTTGTCCA AACTGGAACA ACACTCAACC	5820
CTATCTCGGG CTATTCTTT GATTTATAAG GGATTTGCC GATTTCGGAA CCACCATCAA	5880
ACAGGATTTC CGCCTGCTGG GGCAAACCGAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG	5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCT	6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCAATTAA TGCAGCTGGC	6060
ACGACAGGTT TCCCAGTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC	6120
TCACTCATTA GGCACCCAG GCTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA	6180
TTGTGAGCGG ATAACAATTTCACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC	6240
TACGGCAGCC GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCG AGCTCTTCCC	6300
GCCATCTGAT GAGCAGTTGA AATCTGGAAC TGCCTCTGTT GTGTGCCTGC TGAATAACTT	6360
C3 Cont.	
CTATCCCAGA GAGGCCAAAG TACAGTGGAA GGTGGATAAC GCCCTCCAAT CGGGTAACCTC	6420
CCAGGAGAGT GTCACAGAGC AGGACAGCAA GGACAGCACC TACAGCCTCA GCAGCACCC	6480
GACGCTGAGC AAAGCAGACT ACGAGAAACA CAAAGTCTAC GCCTGCGAAG TCACCCATCA	6540
GGGCCTGAGC TCGCCCGTCA CAAAGAGCTT CAACAGGGGA GAGTGTCTA GAACGCGTCA	6600
CTTGGCACTG GCCGTCGTTT TACAACGTG TGACTGGAA AACCTGGCG TTACCCAAGC	6660
TTTGTACATG GAGAAAATAA AGTGAACAA AGCACTATTG CACTGGCACT CTTACCGTTA	6720
CTGTTACCC CTGTGGCAAA AGCCGCCTCC ACCAAGGGCC CATCGGTCTT CCCCCGGCA	6780
CCCTCCTCCA AGAGCACCTC TGGGGCACA GCGGCCCTGG GCTGCCTGGT CAAGACTAAT	6840
TCCCCGAACC GGTGACGGTG TCGTGGAACT CAGGCGCCCT GACCAGCGGC GTGCACACCT	6900
TCCCGGCTGT CCTACAGTCC TCAGGACTCT ACTCCCTCAG CAGCGTGGTG ACCGTGCCCT	6960
CCAGCAGCTT GGGCACCCAG ACCTACATCT GCAACGTGAA TCACAAGCCC AGCAACACCA	7020
AGGTGGACAA GAAAGCAGAG CCCAAATCTT GTACTAGTGG ATCCTACCCG TACGACGTT	7080
CGGACTACGC TTCTTAGGCT GAAGGCGATG ACCCTGCTAA GGCTGCATTC AATAGTTAC	7140
AGGCAAGTGC TACTGAGTAC ATTGGCTACG CTTGGCTAT GGTAGTAGTT ATAGTTGGTG	7200

CTACCATAGG GATTAATTAA TTCAAAAAGT TTACGAGCAA GGCTTCTTAA GCAATAGCGA 7260
 AGAGGCCCGC ACCGATCGCC CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGCGCTT 7320
 TGCCTGGTTT CCGGCACCAG AAGCGGTGCC GGAAAGCTGG CTGGAGTGCAG ATCTTCCTGA 7380
 GGCCGATACTG GTCGTCGTCC CCTCAAACCTG GCAGATGCAC GGTTACGATG CGCCCACATCTA 7440
 CACCAACGTA ACCTATCCA TTACGGTCAA TCCGCCGTTT GTTCCCACGG AGAATCCGAC 7500
 GGGTTGTTAC TCGCTCACAT TTAATGTTGA TGAAAGCTGG CTACAGGAAG GCCAGACGCG 7560
 AATTATTTT GATGGCGTTC CTATTGGTTA AAAAATGAGC TGATTTAAC AAAATTTAAC 7620
 GCGAATTAA ACAAAATATT AACGTTACA ATTAAATAT TTGCTTATAC AATCTTCCTG 7680
 TTTTGCCCCC TTTTCTGATT ATCAACCGGG GTACATATGA TTGACATGCT AGTTTACGA 7740
 TTACCGTTCA TCGATTCTCT TGTTGCTCC AGACTCTCAG GCAATGACCT GATAGCCTTT 7800
 GTAGATCTCT CAAAAATAGC TACCCCTCTCC GGCATTAATT TATCAGCTAG AACGGTTGAA 7860
 TATCATATTG ATGGTGATTT GACTGTCTCC GGCCTTTCTC ACCCTTTGA ATCTTTACCT 7920
 ACACATTACT CAGGCATTGC ATTAAAATA TATGAGGGTT CTAAAAATTG TTATCCTTGC 7980
 GTTGAAATAA AGGCTTCTCC CGCAAAAGTA TTACAGGGTC ATAATGTTT TGGTACAACC 8040
 GATTTAGCTT TATGCTCTGA GGCTTTATTG CTTAATTTTG CTAATTCTTT GCCTTGCCTG 8100
 TATGATTTAT TGGACGTT 8118

C3
cont.

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(5, "")
- (D) OTHER INFORMATION: /note= "S REPRESENTS EQUAL MIXTURE OF G AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(6, "")
- (D) OTHER INFORMATION: /note= "M REPRESENTS EQUAL MIXTURE OF A AND C"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(8, "")

(D) OTHER INFORMATION: /note= "R REPRESENTS EQUAL MIXTURE OF A AND G"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(11, "")
- (D) OTHER INFORMATION: /note= "K REPRESENTS EQUAL MIXTURE OF G AND T"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(20, "")
- (D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGGTSMARCT KCTCGAGTCW GG

22

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

C3
cont.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGGTCCAGCT GCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGTCCAGCT GCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGTCCAGCT TCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGGTCCAGCT TCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

C³
cont.
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGGTCCAAGT GCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGGTCCAAGT GCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGGTCCAAC TCTCGAGTCT GG

22

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGGTCCAAC TCTCGAGTCA GG

22

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

C3
cont.

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(5..6, "")
- (D) OTHER INFORMATION: /note= "N=INOSINE"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(8, "")
- (D) OTHER INFORMATION: /note= "N=INOSINE"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(11, "")
- (D) OTHER INFORMATION: /note= "N=INOSINE"

(ix) FEATURE:

- (A) NAME/KEY: misc_difference
- (B) LOCATION: replace(20, "")
- (D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AGGTNNANCT NCTCGAGTCW GG

22

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTATTAAC TA GTAACGGTAA CAGTGGTGCC TTGCCCCA

38

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGCTTACTA GTACAATCCC TGGGCACAAT

30

(2) INFORMATION FOR SEQ ID NO:18:

*C³
cont.*

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAGTTCCGA GCTCGTTGTG ACTCAGGAAT CT

32

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCAGTTCCGA GCTCGTGTG ACGCAGCCGC CC

32

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCAGTTCCGA GCTCGTGCTC ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CCAGTTCCGA GCTCCAGATG ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:22:

C3
WNT

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCAGATGTGA GCTCGTGATG ACCCAGACTC CA

32

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCAGATGTGA GCTCGTCATG ACCCAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCAGTTCCGA GCTCGTGATG ACACAGTCTC CA

32

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCAGCATTCT AGAGTTTCAG CTCCAGCTTG CC

32

(2) INFORMATION FOR SEQ ID NO:26:

*C3
cont.*

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCGCCGTCTA GAATTAACAC TCATTCCTGT TGAA

34

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GATCCTAGGC TGAAGGCGAT GACCCTGCTA AGGCTGC

37

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATTCAATAGT TTACAGGCAA GTGCTACTGA GTACA

35

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTGGCTACGC TTGGCTATG GTAGTAGTTA TAGTT

35

(2) INFORMATION FOR SEQ ID NO:30:

C³
cont.

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGTGCTACCA TAGGGATTAA ATTATTCAAA AAGTT

35

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TACGAGCAAG GCTTCCTTA

18

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AGCTTAAGAA GCCTTGCTCG TAAACTTTT GAATAATT

39

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AATCCCTATG GTAGCACCAA CTATAACTAC TACCAT

36

(2) INFORMATION FOR SEQ ID NO:34:

C³
cont.

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

AGCCCAAGCG TAGCCAATGT ACTCAGTAGC ACTTG

35

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CCTGTAAACT ATTGAATGCA GCCTTAGCAG GGTC

34

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ATCGCCTTCA GCCTAG

16

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CATTTTGCA GATGGCTTAG A

21

(2) INFORMATION FOR SEQ ID NO:38:

c3

uny.

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TAGCATTAAC GTCCAATA

18

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ATATATTTA GTAAGCTTCA TCTTCT

26

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GACAAAGAAC GCGTGAAAAC TTT

23

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GCGGGCCTCT TCGCTATTGC TTAAGAAGCC TTGCT

35

(2) INFORMATION FOR SEQ ID NO:42:

*C3
Cont.*

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAACGACGGC CAGTGCCAAG TGACGCGTGT GAAATTGTTA TCC

43

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGCGAAAGGG AATTCTGCAA GGCGATTAAG CTTGGGTAAC GCC

43

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGCGTTACCC AAGCTTGTA CATGGAGAAA ATAAAG

36

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TGAAACAAAG CACTATTGCA CTGGCACTCT TACCGTTACC GT

42

(2) INFORMATION FOR SEQ ID NO:46:

*C³
cont.*

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TACTGTTTAC CCCTGTGACA AAAGCCGCC AGGTCCAGCT GC

42

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TCGAGTCAGG CCTATTGTGC CCAGGGATTG TACTAGTGGA TCCG

44

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGGCGAAAGG GAATTCGGAT CCACTAGTAC AATCCCTG

38

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GGCACAAATAG GCCTGACTCG AGCAGCTGGA CCAGGGCGGC TT

42

(2) INFORMATION FOR SEQ ID NO:50:

*C3
cont.*

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

TTGTCACAGG GGTAAACAGT AACGGTAACG GTAAGTGTGC CA

42

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTGCAATAGT GCTTTGTTTC ACTTTATTTT CTCCATGTAC AA

42

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

TAACGGTAAG AGTGCAGTG C

21

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CACCTTCATG AATTGGCAA GGAGACAGTC AT

32

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

*C3
cont.*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AATTGGCAA GGAGACAGTC AT

22

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATTGTT

39

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

ATTACTCGCT GCCCAACCAG CCATGGCCGA GCTCGTGAT

39

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GACCCAGACT CCAGATATCC AACAGGAATG AGTGTAAAT

39

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 13 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCTAGAACGC GTC

13

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTCAGGTTGA AGCTTACGCG TTCTAGAATT AACACTCATT CCTGT

45

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TGGATATCTG GAGTCTGGGT CATCACGAGC TCGGCCATG

39

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GCTGGTTGGG CAGCGAGTAA TAACAATCCA GCGGCTGCC

39

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

C3
cont.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GTAGGCAATA GGTATTCAT TATGACTGTC CTTGGCG

37

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TGACTGTCTC CTTGGCGTGT GAAATTGTTA

30

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TAACACTCAT TCCGGATGGA ATTCTGGAGT CTGGGT

36

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GCCAGTGCCA AGTGACGCGT TCTA

24

(2) INFORMATION FOR SEQ ID NO:66:

*C3
cont.*

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATATATTTTA GTAAGCTTCA TCTTCT

26

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GACAAAGAAC GCGTGAAAAC TTT

23

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CTGAACCTGT CTGGGACCA	AGTTGATGCT ATAGGATCAG ATCTAGAATT CATTAGAGA	60
----------------------	--	----

CTGGCCTGGC TTCTGC		76
-------------------	--	----

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TCGACCGTTG GTAGGAATAA TGCAATTAAT GGAGTAGCTC TAAATTCAGA ATTCA	TCTAC	60
--	-------	----

<i>C³</i> cont.	ACCCAGTGCA TCCAGTAGCT	80
-------------------------------	-----------------------	----

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GGTAAACAGT AACGGTAAGA GTGCCAG		27
-------------------------------	--	----

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 54 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CGCCTTCAGC CTAAGAAGCG TAGTCCGGAA CGTCGTACGG GTAGGATCCA CTAG

54

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CACCGGTTCG GGGATTAGT CTTGACCAGG CAGCCCAGGG C

41

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 51 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

C₃
cont. (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ATTCCACACA TTATACGAGC CGGAAGCATA AAGTGTCAAG CCTGGGGTGC C

51

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CTGCTCATCA GATGGCGGGA AGAGCTCGGC CATGGCTGGT TG

42

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

GAACAGAGTG ACCGAGGGGG CGAGCTCGGC CATGGCTGGT TG

42

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GGGCTTTTGC CACAGGGGT

19

C³
end.